Unraveling NSA's TURBULENCE Programs

15 September 2014 *by* Robert Sesek

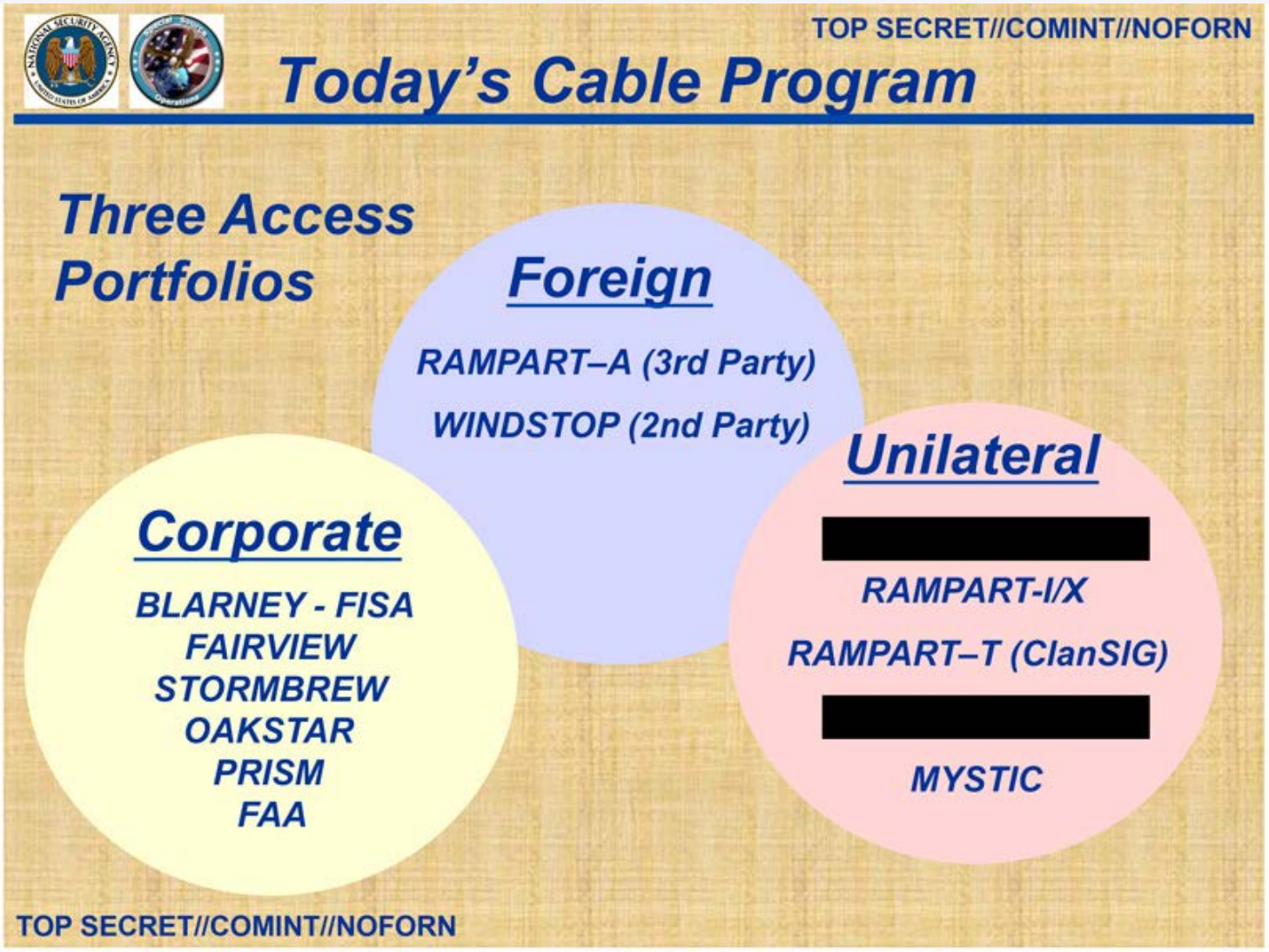
The NSA TURBULENCE program was first revealed in 2007 by the *Baltimore Sun* for being over-budget and poorly-managed [1][2]. Many documents from the Snowden trove from the past two years reference parts of this program, but no coherent story or overview of how these programs operate together has been described. This is an attempt to do so, using leaked and declassified documents.

At the time the *Sun* reported on TURBULENCE in 2007, it had an annual cost approaching $500 million dollars, and it was comprised of nine smaller programs [1]; this article covers five of those that have been revealed to date. The predecessor program to TURBULENCE was called TRAILBLAZER, and it was shut down for cost overruns and mismanagement [3]. The TURBULENCE leadership chose to use “several smaller programs [as] a way to hedge bets on uncertain technology” [2], unlike TRAILBLAZER which was a monolithic project. While TURBULENCE was reported to have had a rocky start, it is now clearly operational and central to NSA’s 21st century SIGINT mission.

Passive Collection: TURMOIL

The TURMOIL program is NSA’s global passive SIGINT apparatus. TURMOIL is a “high-speed passive collection systems intercept [for] foreign target satellite, microwave, and cable communications as they transit the globe” [4]. Presumably all this collection is related to Internet data (DNI), rather than other forms of SIGINT.

NSA collects using TURMOIL through a variety of sources. One of those is the RAMPART-A program, which provides foreign third-party “collection against long-haul international leased communications through special access initiatives with world-wide SIGINT partnerships” [5]. The Internet links that are tapped through RAM-A have “access to over 3 Terabits per second” of data, and “every country code in the world is seen at one or more RAMPART-A collection accesses” [5]. In other words, friendly communications companies provide access to their backbone links, and those “partners work the fiber projects under the cover of an overt Comsat [communications satellite] effort” [6].



Source: Kristian Jensen [7]

It is likely that the other RAMPART projects, such as "RAMPART-M for access to undersea cables" and "RAMPART-T for land-based cables, in cooperation with CIA" [9] contribute to the TURMOIL collection platform. One other specific program that contributes to TURMOIL is the MUSCULAR program, in which the NSA tapped the fiber optic links of Google and Yahoo's datacenters [10]. MUSCULAR and another cable-tapping program called INCENSER are part of the WINDSTOP project, rather than RAMPART, access to which occurs through a second party, reported to be GCHQ [11]. There are likely several other cable-tapping programs that contribute to TURMOIL.

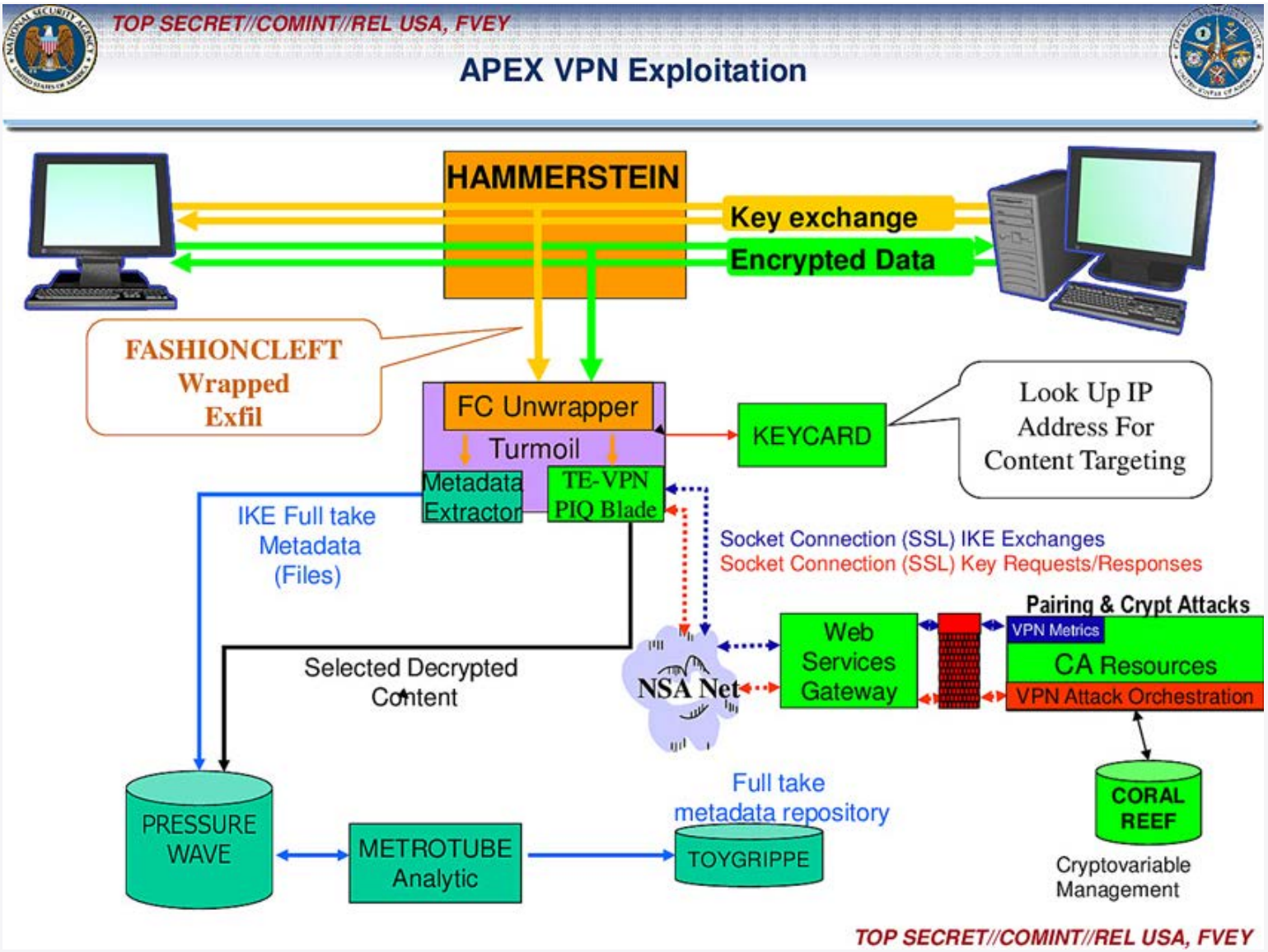
A unique aspect of the TURMOIL program is that it operates on the packet level, rather than operating on full sessions [12]. Communications on the Internet, such as an email or webpage, are broken down into tiny pieces called packets when they are transmitted, allowing each small piece of data to take the fastest route possible to its destination. When

all the packets arrive at the destination, they are reassembled and made whole again. This reassembly, and grouping related email messages and webpages, is what the NSA refers to as “sessionized” data, since the data are full sessions rather than broken-down packets.

Plugin Architecture

Because TURMOIL operates on the packet level, it can choose to handle certain types of traffic with special processing. From leaked documents, it is clear that there are at least three unique types of traffic that are handled at the TURMOIL level.

The first two are part of a system called APEX [13] and handle VPN and VoIP traffic. In the case of VPN traffic, a system called HAMMERSTEIN identifies the traffic and sends the metadata to a database called TOYGRIPPE. The TOYGRIPPE database is a “repository of VPN endpoints”[14] that is used by targeting officers to determine if that computer should be a target for further exploitation [13]. The TURMOIL VPN module also looks up the IP information in a database called KEYCARD to determine if the target should be tasked for targeted SIGINT collection or to recover the VPN key. Of special note is that this VPN traffic passes through a system called “TE-VPN PIQ Blade.” PIQ almost certainly refers to the PICARESQUE ECI marking, which is associated with BULLRUN [16]. The BULLRUN program is NSA’s effort to weaken and exploit encryption that protects digial SIGINT, whether by finding bugs in cryptographic algorithms or by manipulating standards bodies or companies into weakening encryption tools. It is safe to assume that PIQ is a compartment that contains the details of a cryptologic attack against a specific VPN technology (or technologies), which the NSA/GCHQ either found or paid for.

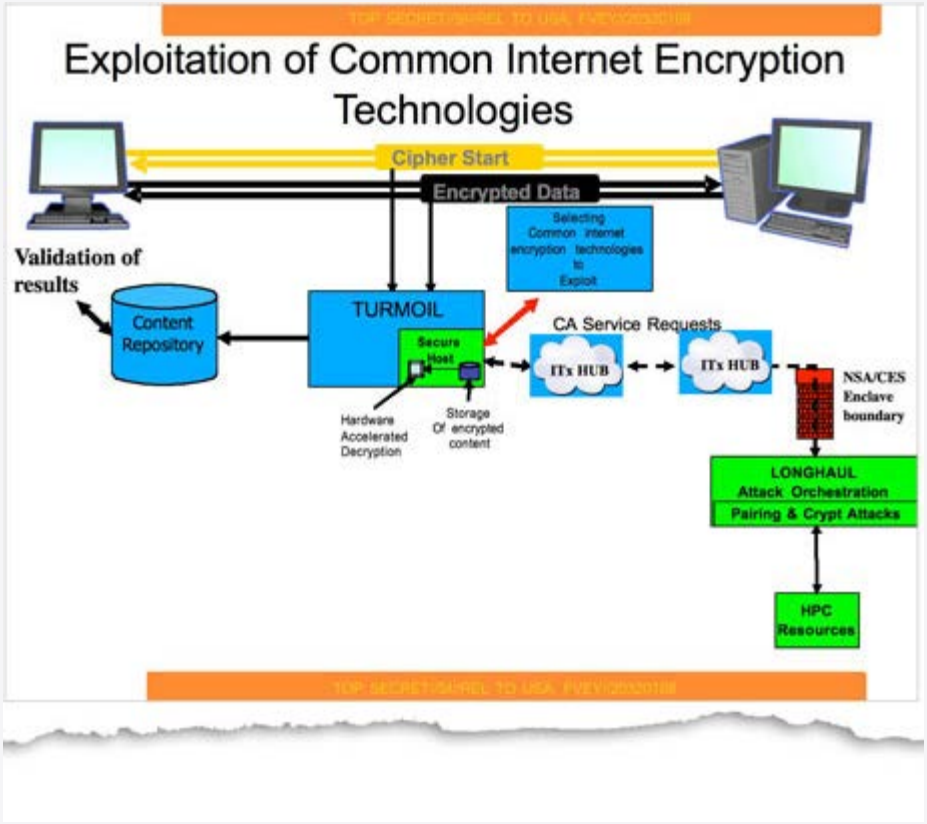


Source: The Intercept [13]

VoIP traffic is processed through a module in TURMOIL called HAMMERCHANT, which processes the digital telephony through CONVEYANCE, for storage in NUCLEON – just like a regular telephone call [13]. In this way, NSA analysts can listen on audio communications regardless of whether it was carried over the PSTN or VoIP.

The third module is the apparent ability to do decryption of encrypted Internet traffic [17]. Only one slide shows TURMOIL interacting with a hardware-accelerated decryption element, but this is likely performing attacks against weak Internet cryptographic standards. This information broke in the stories about the BULLRUN cryptographic exploitation programs, but no strong connection was made between BULLRUN and TURMOIL in previous reporting about these programs. It is unclear if NSA can decrypt (weakly enciphered) Internet traffic in real-time, but based on the diagram's similarities to HAMMERCHANT and HAMMERSTEIN, that

seems possible.



Source: The Guardian [17]

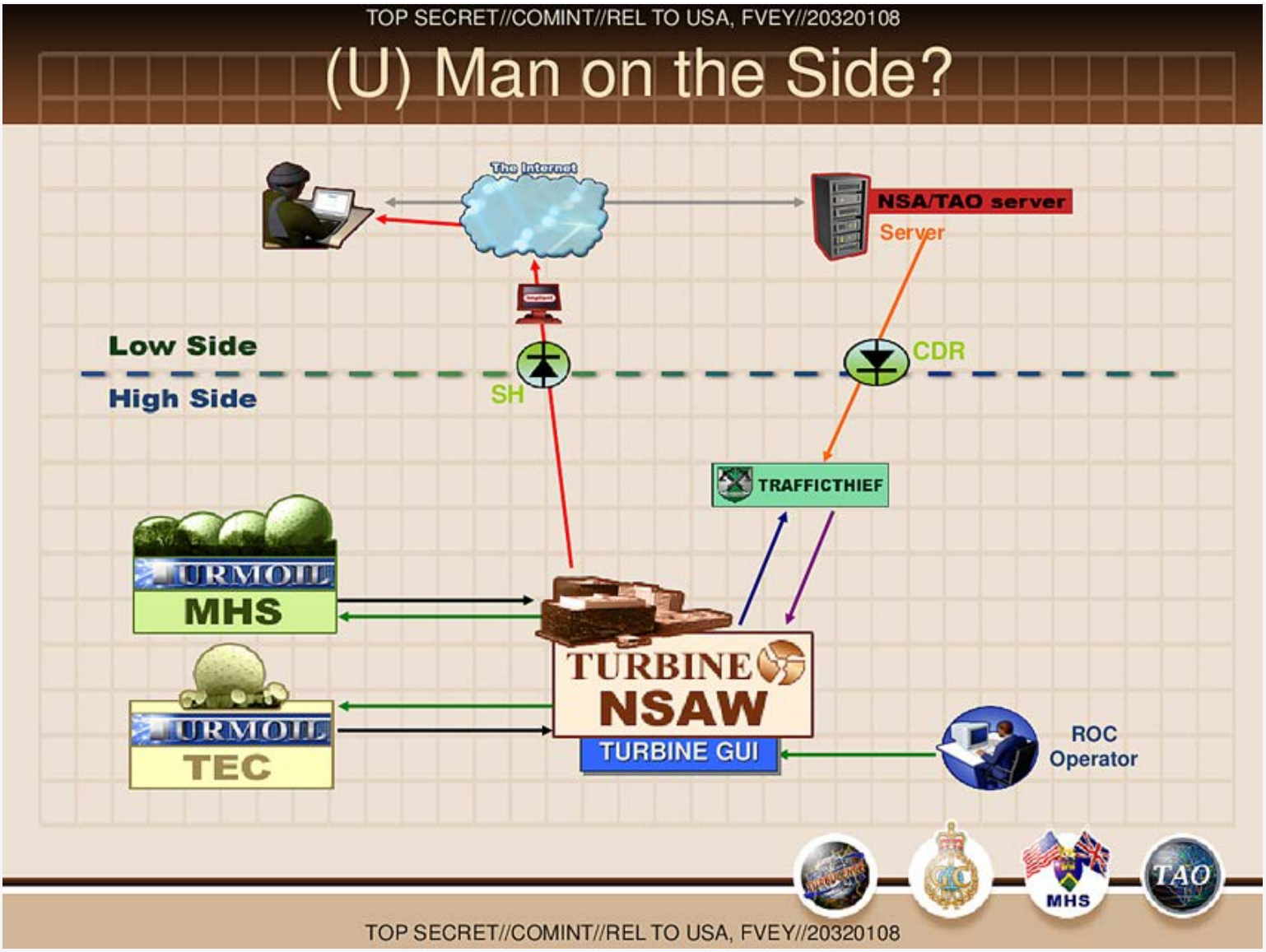
Active Collection: TURBINE

The TURMOIL system sifts through Internet traffic, identifying that which provides useful foreign intelligence information. This system feeds several databases (more on that below), but one of TURMOIL's main counterparts is TURBINE. The TURBINE system is deeply integrated with the QUANTUMTHEORY Computer Network Exploitation (CNE) techniques used by the NSA. When TURMOIL detects packets on the Internet containing a selector (i.e. some identifying information) targeted by NSA, it "tips" the TURBINE system. This "tipping" triggers a program in the TURBINE system to attempt to deploy an active exploit, using the tasked QUANTUMTHEORY attack method, on the target computer. There are several individual QUANTUMTHEORY attack methods from which an officer can select [18]; the end goal for all is to infect the targeted computer such that NSA can maintain access for exfiltrating information.

The TURBINE system "provides centralized automated command/control of a large network of active implants" [21]. Because this system sits between the passive collection system, TURMOIL, and the active attack mechanisms of QUANTUM, it is described as having the

ability to “talk to active and passive sensors/shooters” [24]. One set of slides describes these elements quite succinctly: “Detect: TURMOIL passive sensors detect traffic & tip TURBINE command/control. Decide: TURBINE mission logic constructs response & forwards to TAO [Tailored Access Operations] node” [23]. Once the traffic is forwarded to TAO, the QUANTUMTHEORY attack commences.

QUANTUMTHEORY operates by racing packets across the Internet: when a user attempts to connect to a website, the packets of her request are forwarded across connections until they reach the destination webserver. The webserver then responds with its own stream of packets to deliver the reply. TURMOIL, by virtue of sitting in the middle of the Internet backbone, can detect when a targeted user’s packets are being sent across the Internet. When TURMOIL tips off TURBINE, the QUANTUMINSERT attack attempts to reply to the user’s request by sending its own stream of packets, before the real webserver’s packets can be delivered to the user. If this happens, the packets sent by QUANTUMINSERT quickly redirect the user to a NSA TAO server, which installs the implant. After implantation, the user is redirected back to her original destination, with the hope that the implantation was done covertly. If successful, the NSA can access any data on the computer and maintain this privileged access. (This is why web browser security is serious business).



Source: The Intercept [25]

In this diagram, the components of a QUANTUMINSERT attack are displayed. Two separate TURMOIL systems (each at a collection site) are integrated with the TURBINE system. It is likely that the “MHS” site refers to Menwith Hill Station, one of NSA’s main collection facilities, and “TEC” refers to another, but unknown, site. The TURBINE system also interacts with a database called TRAFFICTHIEF, the purpose of which will be analyzed below. The diagram is also split into two sides: low side and high side, referring to the network on which the part runs. The high side components run in NSA’s classified TOP SECRET NSANet network, to protect itself from discovery (CNE is covert action) and to protect the components from outsiders. The low side components are responsible for implantation, and are managed by TAO on the public Internet. To communicate across the low-high side divide, NSA uses something called a “data diode.” Two are seen in this diagram as green circles with an arrow, one allowing traffic to go from high-to-low and the other for low-to-

high. The diode labeled SH is an abbreviation for the codename SURPLUSHANGAR [22]; it is unknown what the label CDR is an abbreviation for.

According to NSA documents, TURBINE is quite effective and has implanted thousands of computers. However an issue appears to have lessened its efficacy, due to the way the system was designed. On all diagrams, TURMOIL is spread across the globe, but TURBINE is centralized at NSA's headquarters in Maryland (also known as NSAW, for NSA-Washington). When attempting to QUANTUMINSERT a target that is geographically far away from the TURBINE system, the latency caused by communicating between the implantation system, TURBINE, and TURMOIL took too long and QUANTUM was losing the packet race. To work around this problem, NSA came up with QFIRE, which has the goal to "eliminate trans-Atlantic/Pacific latency" [19] by moving the mission logic and decisions to be local to the collection site. With this revised design, TURMOIL no longer has to tip TURBINE back at NSAW; the TURBINE logic resides in the same geographic location as TURMOIL. NSA employees probably still use the same TURBINE interface (GUI) to select, target, and task implantation, but the actual mission logic is now replicated across the globe to each collection site that is QUANTUM-enabled [19].

Integration with XKEYSCORE

The XKEYSCORE program was widely reported, since it is NSA's and GCHQ's rolling buffer of three days worth of full-take Internet traffic [26]. According to a GCHQ document, the design of XKS has evolved to meet the growing traffic demand. The first iteration of the system, called "Traditional" used another system called WEALTHYCLUSTER, which "sessionized all the data on the link and presented it all to XKS" [27]. For higher data rates though, collecting everything was too much, so TURMOIL is used to select interesting packets.

The next design iteration of XKS, called "Stage 2", leveraged TURMOIL's access to terabytes of information flowing over global communication networks. Remember that TURMOIL operates only on packets, not full sessions. To feed XKEYSCORE, it only collects 5% of those packets, based on the data meeting various criteria including strong selection (e.g. an email or keyword), subnet promotion (specific networks), technology promotion (e.g. Tor), and email domains [27]. Once TURMOIL has collected and saved the selected packets, it forwards them to XKEYSCORE, which sessionizes the data and makes it searchable to NSA analysts through the XKS interface.

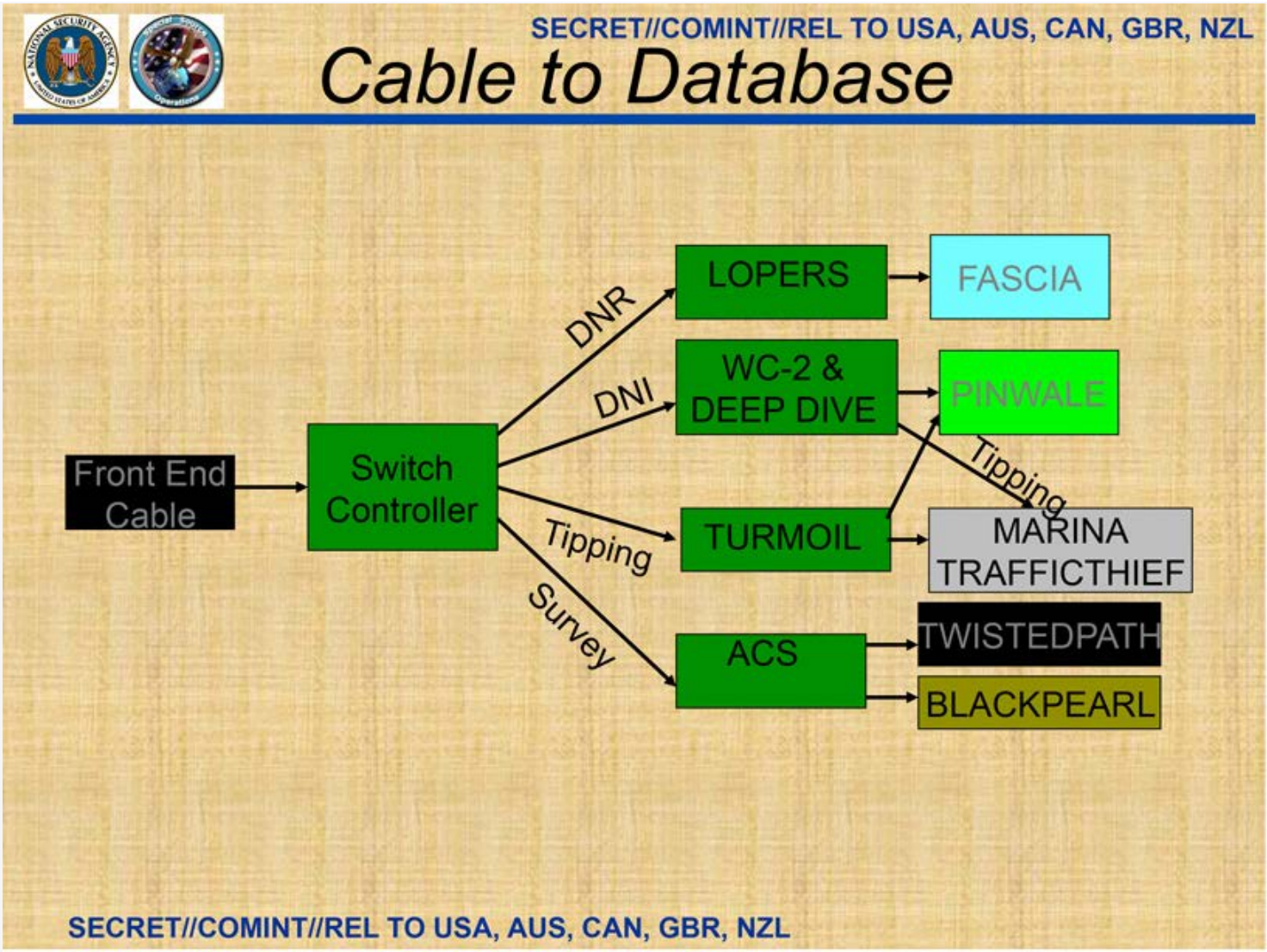
The most advanced design iteration is called “Deep Dive”, and it can sessionize at “10G data rates” [27]. To do this, all data are sessionized on the link, and the interesting data are “promoted” (that is, selected for storage) using the GENESIS language. It is likely that the XKEYSCORE source code that Jacob Applebaum published, demonstrating promotion of Tor traffic, is written in this GENESIS language [28].

Data Storage and TRAFFICTHIEF

Most Internet metadata are stored in NSA’s main database for that information, MARINA. However a separate database operated under TURBULENCE called TRAFFICTHIEF also exists and integrates with XKEYSCORE [29]. TRAFFICTHIEF is described as containing “meta-data from a subset of tasked strong-selectors” [30]. On various diagrams, TRAFFICTHIEF appears to primarily be used by TURBINE and TAO’s implantation servers [25]. I estimate that this database contains the selectors used to target computers and individuals for TAO operations, QUANTUMTHEORY and otherwise. It is likely that the reason TRAFFICTHIEF also feeds into XKEYSCORE is because if a target is important enough to collect on with active SIGINT, then analysts will want data about the target from all SIGINT databases the NSA has, both passive and active.

Cross-Feeding Databases

As part of cable-tapping programs like RAMPART, WINDSTOP, and Upstream, the NSA takes the signal from the fiber optic cable and sends it through several different systems:



Source: Ryan Gallagher [8]

The “Front-End Cable” in this case is part of the communication provider’s network. The “Switch Controller” is the first piece of NSA’s infrastructure, and it is likely a cable splitter. I analyze below how this may be a project codenamed TUMULT. It is apparent from this diagram that Internet data are sent through two processing systems. In the diagram, “Tipping” data are shown to be sent to TURMOIL, as discussed above, which sends metadata to both MARINA (long-term metadata repository) and TRAFFICTHIEF (TURBINE tipping metadata repository). It is also shown from this diagram that TURMOIL sends full content to the PINWALE database for storage.

The diagram also shows that full DNI data passes through a pair of systems, WC-2 and DEEP DIVE, before sending data to the main content repository PINWALE. It is likely that WC-2 refers to WEALTHYCLUSTER-2, a new version of the system discussed above; and DEEP DIVE refers to the third design iteration of the XKEYSCORE architecture. Since

TURMOIL operates on the packet level, it is more suited to identifying metadata quickly and in real-time. For content, NSA uses WEALTHYCLUSTER to collect packets and sessionize them, which is a logical step for feeding a database that stores full communication content, like PINWALE. It is likely that NSA uses WEALTHYCLUSTER-2 and DEEP DIVE to maintain a majority of traffic displayed in XKEYSCORE (the stated goal of which is to have a three-day buffer of all Internet traffic), with the targeted communications reserved in PINWALE. This contrasts to TURMOIL, the main function of which is to sift through and sort metadata, but it can also feed XKEYSCORE in a “Stage 2” configuration described above.

From the cable taps, NSA also filters out telephony data using LOPERS, which is a software-based system for processing PSTN data [31]. The data extracted by LOPERS is stored in the FASCIA database, which is NSA’s repository of cell phone metadata (including location, IMEI, IMSI, and other unique identifiers) [32].

It is unclear to what “ACS” refers in the last prong of the processing diagram, but the data goes to two separate databases: TWISTEDPATH and BLACKPEARL. It is unknown what data are stored in TWISTEDPATH, but BLACKPEARL is used to store “SIGINT session 5-tupel, identified routers, routing protocols, SIGINT access points (inferred SIGINT access points)” [15]. The BLACKPEARL system is associated with TREASUREMAP, a program to map the Internet [15]. Since this is the “Survey” prong of the processing diagram, as in survey and map the Internet, it is likely that TWISTEDPATH stores related data.

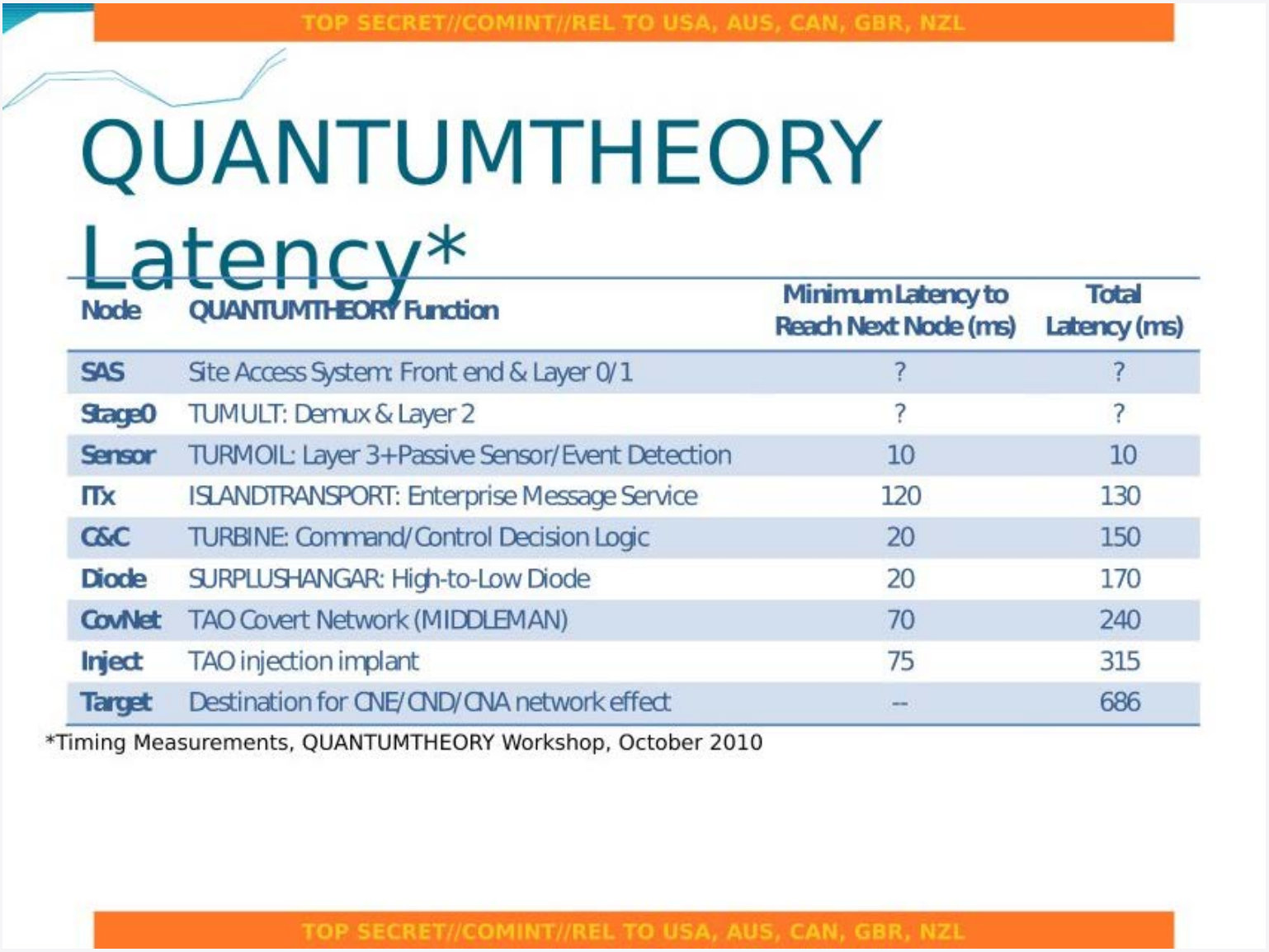
Active Defense: TUTELAGE

This component of TURBULENCE is unlike the others in that it faces inward. The goal of the TUTELAGE program is to defend Department of Defense networks [20]. The *Wall Street Journal* reports that the program “detects incoming cyberattacks and allows NSA to block the threat or manipulate the attack code” [33]. Similarly, the *Washington Post* reports that the system “has the ability to decide how to handle malicious intrusions – to block them or watch them closely to better assess the threat” [34]. It has also been associated with the BYZANTINE-HADES program, which refers to the “concerted effort against Chinese hackers. It now has a new name. Probably containing the word ‘LEGION’” [35]. The TUTELAGE system has been re-purposed to protect civilian networks, like the DHS, under the name EINSTEIN-3 [36]. Little more is known about the details of this program. Media outlets likely did not

publish any documents Snowden had about this program, because doing so could inflict actual, articulable harm against the U.S. Government, if its network defense strategy were made public.

Small Mention: TUMULT

Another project under TURBULENCE is TUMULT. It is only described on one slide, in relation to QUANTUMTHEORY and TURMOIL. It is referenced in a slide describing the latencies of various nodes in the QUANTUMTHEORY stack:



TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

QUANTUMTHEORY

Latency*

Node	QUANTUMTHEORY Function	Minimum Latency to Reach Next Node (ms)	Total Latency (ms)
SAS	Site Access System: Front end & Layer 0/1	?	?
Stage0	TUMULT: Demux & Layer 2	?	?
Sensor	TURMOIL: Layer 3+Passive Sensor/Event Detection	10	10
ITx	ISLANDTRANSPORT: Enterprise Message Service	120	130
C&C	TURBINE: Command/Control Decision Logic	20	150
Diode	SURPLUSHANGAR: High-to-Low Diode	20	170
CovNet	TAO Covert Network (MIDDLEMAN)	70	240
Inject	TAO injection implant	75	315
Target	Destination for CNE/CND/CNA network effect	--	686

*Timing Measurements, QUANTUMTHEORY Workshop, October 2010

TOP SECRET//COMINT//REL TO USA, AUS, CAN, GBR, NZL

Source: Der Spiegel [22]

The first node is SAS, described as the “Site Access System Frontend & Layer 0/1”. Stacked on top of that is TUMULT, described only as “Demux & Layer 2”. Above this layer is TURMOIL, described on the slide as “Layer 3 + passive sensor/event detection”. If the layers

refer to the OSI 7 Layer Model, which I think is likely, then the SAS frontend deals with the actual physical fiber optic cable and is a hardware component. The TUMULT component is also likely a hardware module, and if it operates at layer two of OSI, then it refers to the data link layer. It is possible that this is a piece of physical hardware that performs the actual traffic splitting of the fiber cable, one part of the split transiting as it should, the other being sent to TURMOIL for processing.

The rest of the latency table describes the full flow of QUANTUMTHEORY. After the TURMOIL sensor detects a target, it uses ISLANDTRANSPORT, an “Enterprise Message Service” to communicate with TURBINE. This is the tipping mentioned above, and ISLANDTRANSPORT is likely something similar to Protocol Buffers RPC or Thrift. TURBINE makes the decision on whether to implant, probably in consultation with TRAFFICTHIEF as discussed above. The request for implantation then crosses the SURPLUSHANGAR diode to traverse the low-side/high-side boundary. The covert network that TAO runs is called MIDDLEMAN, and it then performs the packet injection to implant the computer program on the target’s computer. According to the NSA slide, QUANTUMTHEORY can implant a target in 686 milliseconds.

Closing Remarks

In all the documents about the TURBULENCE program, no mention is made of what authorities the projects operate under. The NSA does not perform collection unless it believes it has the legal authority to do so; programs like PRISM operate under FAA authority, whereas almost all foreign intelligence is collected under EO 12333. It is common in many leaked briefing presentations to state the authority under which collection occurs. But such information is absent in any TURBULENCE presentations. Given the slide entitled “SSO [Special Source Operations]: Today’s Cable Program”[7], I believe that TURBULENCE, specifically TURMOIL, is a collection platform for **all** kinds of NSA access. The “three access portfolios” displayed on the slide cover domestic cooperation, foreign cooperation, and “unilateral” access. From a resources management perspective, it would only make sense for NSA to build one large, all-encompassing SIGINT platform, TURBULENCE and TURMOIL, and store the data with special labels based on the authority under which it was collected. There are repeated references to storing FISA data with labels such as RAGTIME [37]. And being able to reuse the same platforms and systems would simplify analyst access, pool software development resources, and unify systems administration tasks. I believe the TURBULENCE

project was NSA's program to modernize its SIGINT systems for the 21st century, as communication switched from radio and telegraph to the Internet. It appears to be wildly successful.

This page was last revised 15 September 2014; if new information comes out that either further corroborates or disproves this analysis, this page will be updated.

Footnotes

1. Siobhan Gorman, "NSA program draws Congress' ire", *Baltimore Sun*, 28 March 2007. Accessed 13 September 2014.
2. Siobhan Gorman, "Costly NSA initiative has a shaky takeoff", *Baltimore Sun*, 11 February 2007. Accessed 13 September 2014.
3. Siobhan Gorman, "System Error", *The Baltimore Sun*, 29 Jaunuary 2006. Accessed 14 September 2014.
4. "Turbine and Turmoil", *The Intercept*, 12 March 2014. Accessed 13 September 2014.
5. Kristian Jensen, "Black Budget", *Information*, 19 June 2014. Accessed 13 September 2014.
6. Kristian Jensen, "RAMPART-A Project Overview", page 10, *Information*, 19 June 2014. Accessed 13 September 2014.
7. Kristian Jensen, "Special Source Operations", page 2, *Information*, 19 June 2014. Accessed 13 September 2014.
8. Ibid., page 8.
9. "The National Security Agency in 2002", *Top Level Communications*, 3 July 2014. Accessed 13 September 2014.
10. Barton Gellman, Ashkan Soltani, and Andrea Peterson, "How we know the NSA had access to internal Google and Yahoo cloud data", *The Washington Post*, 4 November 2013. Accessed 13 September 2014.
11. "What are SIGADs starting with DS for?", *Top Level Communications*, 15 October 2013. Accessed 13 September 2014.
12. "German, NSA SIGINTers Share DNI Processing Knowledge", *Der Spiegel*, 18 June 2014. Accessed 13 September 2014.
13. "VPN and VoIP Exploitation with HAMMERCHANT and HAMMERSTEIN", *The Intercept*, 12 March 2014. Accessed 13 September 2014.
14. "Bad Guys Are Everywhere", page 16, *Der Spiegel*, 14 September 2014. Accessed 14 September 2014.

15. Ibid., page 15.
16. "Project BULLRUN – classification guide to the NSA's decryption program", page 2, *The Guardian*, 5 September 2013.
17. Glenn Greenwald, "Revealed: how US and UK spy agencies defeat internet privacy and security", *The Guardian*, 5 September 2013. Accessed 13 September 2014.
18. "There is More Than One Way to QUANTUM", *The Intercept*, 12 March 2014. Accessed 13 September 2014.
19. "Getting Close to the Adversary: Forward-based Defense with QFIRE", photo 8, *Der Spiegel*, 30 December 2013. Accessed 13 September 2014.
20. Ibid., photo 5.
21. Ibid., photo 6.
22. Ibid., photo 16.
23. Ibid., photo 7.
24. "Industrial-Scale Exploitation", *The Intercept*, 12 March 2014. Accessed 13 September 2014.
25. "QUANTUMTHEORY", page 4, *The Intercept*, 12 March 2014. Accessed 13 September 2014. Accessed 13 September 2014.
26. Additional XKEYSCORE Slides, Gunnar Rensfeldt, 11 December 2013.
27. "XKeyscoreTabs XKS Development", *Der Spiegel*, 18 June 2014. Accessed 13 September 2014.
28. von J. Appelbaum, A. Gibson, J. Goetz, V. Kabisch, L. Kampf, and L. Ryge, "NSA targets the privacy-conscious", *Das Erste*, 7 March 2014. Accessed 13 September 2014.
29. "WHAT A WONDERFUL SUCCESS!", *NSA*, Undated. Accessed 13 September 2014.
30. Glenn Greenwald, "XKeyscore: NSA tool collects 'nearly everything a user does on the internet'", *The Guardian*, 31 July 2013. Accessed 13 September 2014.
31. Von Christian Stöcker, "NSA Documentation of Spying in Germany", photo 4, *Der Spiegel*, 29 July 2013. Accessed 13 September 2014.
32. Ashkan Soltani and Matt DeLong, "FASCIA: The NSA's huge trove of location records", *The Washington Post*, 4 December 2013. Accessed 13 September 2014.
33. Siobhan Gorman, "NSA Chief Seeks Bigger Cybersecurity Role", *The Wall Street Journal*, 27 February 2012. Accessed 13 September 2014.
34. Ellen Nakashima, "Cybersecurity Plan to Involve NSA, Telecoms", *The Washington Post*, 3 July 2009. Accessed 13 September 2014.

35. Paulmd199, Tweet, 20 April 2014. Accessed 13 September 2014.
36. Kim Zetter, "NSA Shields Government Networks with AT&T Secret Rooms", *WIRED*, 6 July 2009. Accessed 13 September 2014.
37. Charlie Savage, "Classification Guide for FISA, the Protect America Act and the FISA Amendments Act", page 11, *The New York Times*, 11 March 2014. Accessed 13 September 2014.